

1125-00-1872

Sameeksha Khillan* (sameeksha.khillan@gmail.com), **Edgar Sucar** (edgarsucar@gmail.com), **Xiran Liu** (liu.xiran@wustl.edu) and **Zichao Li** (lizichao7@gmail.com). *Optimizing the Quality-Cost Trade-off of Human Annotation for labeling web-pages to train web page classifiers.*

Google gets human annotators to get training data for binary web page classifiers. Since this data is used as gold standard for training and evaluating classification models, its accuracy, based on the quality of human annotation, is of great importance to the classifier's performance. We want to develop and analyze strategies to mitigate human mistakes to achieve highest possible classifier performance respecting time and budget constraints. We simulate human annotated data based on previous records of rater accuracy and test these strategies. We develop two strategies: the first is an iterative multiple annotation strategy in which multiple humans annotate each web page with a particular distribution of raters (optimized for best data quality) assigned to each web page in advance. The second is based on prior probability of human error combined with confidence output by the classifier to clean evaluation data. We believe these strategies offer solutions to optimize quality-cost trade-off of human annotation. We observe that certain multiple rating orders and combinations lead to better data quality and also data quality is more relevant than data size for better performance, which indicates that a better strategy for rater combination is to spend more on better quality raters. (Received September 19, 2016)